US 20140130158A1

(19) United States

(12) Patent Application Publication

(10) Pub. No.: US 2014/0130158 A1

(43) Pub. Date: May 8, 2014

WANG et al.

(54) **IDENTIFICATION OF MALWARE DETECTION SIGNATURE CANDIDATE CODE**

(71) Applicant: **MICROSOFT CORPORATION**, Redmond, WA (US)

(72) Inventors: **Xun WANG**, Sammamish, WA (US); **Hong JIA**, Bellevue, WA (US)

(73) Assignee: **MICROSOFT CORPORATION**, Redmond, WA (US)

(21) Appl. No.: **13/670,529**

(22) Filed: **Nov. 7, 2012**

Publication Classification

(51) **Int. Cl.**
G06F 21/00 (2006.01)

(52) **U.S. Cl.**
USPC **726/23**

(57) **ABSTRACT**

A region of HTML or PDF file bytecode run on a virtual machine is identified as possible malware, allowing a detection signature to be generated. A determination is made, based on code behavior, that malware may be present. Variables visible in this identification start state can be found by mapping the start state to scopes in an abstract syntax data structure. Searching previously executed states of the virtual machine for any assignment of a variable that belongs to the set of variables of interest provides a set of assignments of interest, even in obfuscated code. Nonterminated assignments of interest will lead in turn to other variables of interest and assignments of interest, until all assignments of interest are terminated. At that point, a region of code defined by the assignments of interest is identified as a malware detection signature generation candidate, and submitted to a human or automated analyst.